1. Affine Spaces, Convex Sets and Simplices

1.1 Note: In this section, we let $\mathbb{F}$ denote a fixed field (for example $\mathbb{F} = \mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{Z}_p$ with p prime) and let W be a fixed vector space over $\mathbb{F}$.

1.2 Definition: An **affine space** in W is a set of the form

$$P = p + U = \{p + u \mid u \in U\}$$

for some element $p \in W$ and some subspace $U \subseteq W$. An element in an affine space is called a **point**.

1.3 Example: Every subspace $U \subseteq W$ is also an affine space in W (since $U = 0 + U$). For an element $a \in U$, we can call the element a a **vector** if we are considering U as a vector space, and we can call the element a **point** if we are considering U as an affine space.

1.4 Example: For a subspace $U \subseteq W$, the **quotient space** W/U is the set

$$W/U = \{p + U \mid p \in W\}.$$

The operations in W/U are given by $(p+U)+(q+U) = (p+q)+U$ and $t(p+U) = tp+U$.

1.5 Theorem: *Let $p, q \in W$ be points and let $U, V \subseteq W$ be subspaces. Then*
(1) $p + U \subseteq q + V$ if and only if $U \subseteq V$ and $p - q \in V$, and
(2) $p + U = q + V$ if and only if $U = V$ and $p - q \in U$.

Proof: Suppose that $p + U \subseteq q + V$. Since $p = p + 0 \in p + U$, we also have $p \in q + V$, say $p = q + v$ where $v \in V$. Then $p - q = v \in V$. Let $u \in U$. Then we have $p + u \in p + U$ and so $p + u \in q + V$, say $p + u = q + w$ where $w \in V$. Then $u = w - (p - q) = w - v \in V$. Conversely, suppose that $U \subseteq V$ and $p - q \in V$, say $p - q = v \in V$. Let $a \in p + U$, say $a = p + u$ where $u \in U$. Then we have $a = p + u = (q + v) + u = q + (u + v) \in q + V$ since $u + v \in V$. This proves part (1), and part (2) follows immediately from part (1).

1.6 Definition: Let P be an affine space in W, say $P = p + U$ where $p \in W$ is a point and $U \subseteq W$ is a subspace. The vector space U, which by the above theorem is uniquely determined, is called the **associated vector space** of P, and we say the P is the affine space through p in the direction of U. We define the **dimension** of P to be

$$\dim(P) = \dim(U).$$

Similarly, the **codimension** of P in W is $\text{codim}_W(P) = \text{codim}_W(U) = \dim(W/U)$.

1.7 Definition: A **line** in W is a 1-dimensional affine space in W. A **plane** in W is a 2-dimensional affine space in W. We often call a 0-dimensional affine space in W a **point** (although, strictly speaking, a 0-dimensional affine space in W is a one-element set which contains a point). A **hyperplane** in W is an affine space in W of codimension 1 (so when $\dim(W) = n$, a hyperplane in W is an $(n-1)$-dimensional affine space in W).

1.8 Example: Let $u_1, u_2, \cdots, u_k \in \mathbb{F}^n$, let $\mathcal{A} = \{u_1, u_2, \cdots, u_k\}$, let $U = \operatorname{Span}\mathcal{A}$, let $p \in \mathbb{F}^n$, $P = p + U$, and let $A = (u_1, u_2, \cdots, u_k) \in M_{n\times k}(\mathbb{F})$. Note that

$$U = \operatorname{Span}\mathcal{A} = \Big\{ \sum_{i=1}^{k} t_i u_i \Big| \text{each } t_i \in \mathbb{F} \Big\} = \big\{ At \,\big|\, t \in \mathbb{F}^k \big\} = \operatorname{Col}(A).$$

We can calculate $\dim(P) = \dim(U)$ in several ways. For example, we can row reduce the matrix A to obtain a reduced row-echelon matrix R. If the pivots in R occur in columns $1 \le j_1 \le j_2 \le \cdots \le j_r \le n$, then $\{u_{j_1}, u_{j_2}, \cdots, u_{j_r}\}$ (the set of corresponding columns in A) is a basis for $U = \operatorname{Col}(A)$ and we have $\dim(P) = \dim(U) = r = \operatorname{rank}(A)$. Alternatively, we can row-reduce the matrix A^T to obtain a row-reduced echelon matrix S. The nonzero rows of S then form a basis for $\operatorname{Row}(S) = \operatorname{Row}(A^T) = U$.

1.9 Example: Let $A \in M_{k\times n}(\mathbb{F})$ and let $b \in \mathbb{F}^k$. If P is the solution set

$$P = \big\{ x \in \mathbb{F}^n \,\big|\, Ax = b \big\}$$

then either $P = \emptyset$ (the empty set) or P is an affine space in $\mathbb{F}^n$. Indeed if $p \in \mathbb{F}^n$ is in the solution set so that $Ap = b$, then for $x \in \mathbb{F}^n$ we have

$$Ax = b \iff Ax = Ap \iff A(x-p) = 0 \iff (x-p) \in \operatorname{Null}(A) \iff x \in p + \operatorname{Null}(A)$$

and so the solution set is the affine space $P = p + U$ where $U = \operatorname{Null}(A)$. We can determine whether $Ax = b$ has a solution, and if so we can determine a solution and find a basis for $U = \operatorname{Null}(A)$ using Gauss-Jordan elimination. We row reduce the augmented matrix $(A|b)$ to obtain a row-reduced augmented matrix, say

$$(A|b) \sim \left(\begin{array}{c|c} R & c \\ 0 & d \end{array} \right)$$

where R is in row reduced echelon form with non-zero rows. If $d \ne 0$ then there is no solution and if $d = 0$ then the solution is obtained from R and c as follows. Let $1 \le j_1 \le j_2 \le \cdots \le j_r \le n$ be the pivot column indices and let $0 \le l_1 \le l_2 \le \cdots \le l_s \le n$ be the non-pivot column indices in R (so that $r + s = n$). Let $u_1, u_2, \cdots, u_n$ be the columns of R so $R = (u_1, u_2, \cdots, u_n) \in M_{r\times n}(\mathbb{F})$. Write $R_J = (u_{j_1}, u_{j_2}, \cdots, u_{j_r}) \in M_{r\times r}(\mathbb{F})$ and $R_L = (u_{l_1}, u_{l_2}, \cdots, u_{l_s}) \in M_{r\times s}(\mathbb{F})$ and note that $R_J = I$. For $B \in M_{n\times s}(\mathbb{F})$ with row vectors $v_1, v_2, \cdots, v_n \in \mathbb{F}^s$ so that we have $B = (v_1, \cdots, v_n)^T \in M_{n\times s}(\mathbb{F})$, write $B_J = (v_{j_1}, v_{j_2}, \cdots, v_{j_r})^T \in M_{r\times s}(\mathbb{F})$ and $B_L = (v_{l_1}, v_{l_2}, \cdots, v_{l_s})^T \in M_{s\times s}(\mathbb{F})$, and for $p \in \mathbb{F}^n$ write $p_J = (p_{j_1}, p_{j_2}, \cdots, p_{j_r})^T \in \mathbb{F}^r$ and $p_L = (p_{l_1}, p_{l_2}, \cdots, p_{l_s})^T \in \mathbb{F}^s$. Then the solution to $Ax = b$ is given by

$$x = p + Bt \text{ , where } p_J = c,\ p_L = 0,\ B_J = -R_L \text{ and } B_L = I.$$

Because the matrix $B \in M_{n\times r}(\mathbb{F})$ includes s linearly independent rows, namely the rows in $B_L = I$, it follows that the columns of B are linearly independent and form a basis for $U = \operatorname{Null}(A) = \operatorname{Col}(B)$.

1.10 Example: Let $L : W \to V$ be a linear map and let $a \in W$ and $b \in V$. If $b \notin \operatorname{Range}(L)$ then $L^{-1}(b) = \emptyset$. If $b \in \operatorname{Range}(L)$ with $L(a) = b$ then $L^{-1}(b) = a + \operatorname{Null}(L)$ because for $x \in W$ we have $L(x) = b \iff L(x-a) = L(x) - L(a) = b - b = 0$.

1.11 Theorem: *Let K be a non-empty set and for each $k \in K$, let P_k be an affine space in W. Let $Q = \bigcap\limits_{k\in K} P_k$. Then either $Q = \emptyset$ or Q is an affine space in W.*

Proof: Suppose that $Q \neq \emptyset$. Choose $p \in Q$. For each $k \in K$, let U_k be the associated vector space of P_k, and note that, since $p \in P_k$, we have $P_k = p + U_k$. Let $U = \bigcap_{k\in K} U_k$. Note that U is a subspace of W. Indeed $0 \in U$ and if $u, v \in U$ and $t \in \mathbb{F}$, then for every $k \in K$ we have $u, v \in U_k$ so that $u+v \in U_k$ and $tu \in U_k$, and hence $u+v \in U$ and $tu \in U$. We claim that $Q = p + U$. Let $x \in Q = \bigcap_{k\in K} P_k = \bigcap_{k\in K}(p + U_k)$. For each $k \in K$, choose $u_k \in U_k$ so that $x = p + u_k$. Fix $\ell \in K$ and let $u = u_\ell$. Note that for all $k \in K$ we have $u_k = x - p = u_\ell = u$. Thus $u \in \bigcap_{k\in K} U_k = U$ and we have $x = p + u \in p + U$. Conversely, let $y \in p+U$, say $y = p+u$ with $u \in U$. Then for every $k \in K$ we have $u \in U_k$ so $y = p + u \in p + U_k = P_k$. Since $y \in P_k$ for all k, we have $y \in \bigcap_{k\in K} P_k = Q$.

1.12 Definition: Let $\emptyset \neq \mathcal{A} \subseteq \mathcal{W}$. We define the **affine span** of $\mathcal{A}$, denoted by $\langle \mathcal{A} \rangle$, to be the smallest affine space in W which contains $\mathcal{A}$, or equivalently, the intersection of all affine spaces in W which contain $\mathcal{A}$. Sometimes we omit set brackets from our notation, so for example when $a_0, a_1, \cdots, a_l \in \mathbb{F}^n$ we usually write $\langle a_0, a_1, \cdots, a_l \rangle = \big\langle \{a_0, a_1, \cdots, a_l\} \big\rangle$.

1.13 Theorem: *Let $\emptyset \neq A \subseteq W$ and let $p \in A$. Let $U = \operatorname{Span}\big\{a - p \,\big|\, a \in A\big\}$. Then*

$$\langle A \rangle = p + U = \Big\{ \sum_{i=0}^{n} s_i a_i \,\Big|\, n \in \mathbb{N}\,,\, a_i \in A\,,\, s_i \in \mathbb{F}\,,\, \sum_{i=0}^{n} s_i = 1 \Big\}.$$

Proof: Let $a \in A$. Then $a - p \in U$ and so $a \in p + U$. Thus $p + U$ is an affine space in W which contains A, and so we have $\langle A \rangle \subseteq p + U$. Let Q be any affine space in W which contains A. Note that since $p \in A$, we also have $p \in Q$. Let V be the associated vector space of Q so that $Q = p + V$. For every $a \in A$ we have $a \in Q = p + V$ and so $a - p \in V$. It follows that $U = \operatorname{Span}\big\{a - p \,\big|\, a \in A\big\} \subseteq V$. Since $U \subseteq V$ we have $p + U \subseteq p + V = Q$. Since $p + U \subseteq Q$ for every affine space Q which contains A, it follows that $p + U \subseteq \langle A \rangle$. Thus we have shown that

$$\begin{aligned}\langle A \rangle &= p + U = p + \operatorname{Span}\big\{a - p \,\big|\, a \in A\big\} \\ &= \Big\{ p + \sum_{i=1}^{n} t_i(a_i - p) \Big| n \in \mathbb{N}, a_i \in A, t_i \in \mathbb{F} \Big\}.\end{aligned}$$

Finally note that

$$p + \sum_{i=1}^{n} t_i(a_i - p) = \big(1 - \sum_{i=1}^{n} t_i\big)\, p + \sum_{i=1}^{n} t_i a_i = \sum_{i=0}^{n} s_i a_i$$

where $a_0 = p$, $s_0 = \big(1 - \sum\limits_{i=1}^{n} t_i\big)$, and $s_i = t_i$ for $i \geq 1$.

1.14 Definition: Let $\emptyset \neq A \subseteq W$. An **affine combination** on A is a point in $\langle A \rangle$, that is a point in W of the form

$$p = \sum_{i=0}^{n} s_i a_i \text{ where } n \in \mathbb{N}\,,\, a_i \in A\,,\, s_i \in \mathbb{F}\,,\, \sum_{i=0}^{n} s_i = 1.$$

1.15 Definition: Let $\emptyset \neq A \subseteq W$. We say that A is **affinely independent** when for all $n \in \mathbb{N}$, for all distinct $a_0, a_1, \cdots a_n \in A$ and for all $s_0, s_1, \cdots, s_n \in \mathbb{F}$,

$$\text{if } \sum_{i=0}^{n} s_i a_i = 0 \text{ and } \sum_{i=0}^{n} s_i = 0 \text{ then every } s_i = 0.$$

Otherwise we say that A is **affinely dependent**.

1.16 Remark: Let $\emptyset \neq A \subseteq W$. Then A is affinely independent if and only if every element in $\langle A \rangle$ can be expressed uniquely, up to order, as an affine combination of distinct elements in A.

1.17 Theorem: *Let $\emptyset \neq A \subseteq W$, let $p \in A$, and let $B = \{a - p \,|\, a \in A \setminus \{p\}\}$. Then A is affinely independent if and only if $\mathcal{B}$ is linearly independent.*

Proof: We prove one direction of the if and only if statement and leave the proof of the other direction as an exercise. Suppose that A is affinely independent. Let $n \in \mathbb{Z}^+$, let $u_1, u_2, \cdots, u_n$ be distinct elements in B, let $t_1, \cdots, t_n \in \mathbb{F}$ and suppose that $\sum_{i=1}^{n} t_i u_i = 0$. Note that B is a set of non-zero vectors so each $u_i \neq 0$. Let $a_i = u_i + p$ and note that $a_1, \cdots, a_n$ are distinct elements in A with each $a_i \neq p$. Let $a_0 = p$, let $s_i = t_i$ for $1 \leq i \leq n$ and let $s_0 = -\sum_{i=1}^{n} s_i$. Note that

$$\sum_{i=1}^{n} t_i u_i = 0 \iff \sum_{i=1}^{n} t_i (a_i - p) = 0 \iff \sum_{i=0}^{n} s_i a_i = 0.$$

Since A is affinely independent we have $s_i = 0$ for $0 \leq i \leq n$ and hence $t_i = 0$ for $1 \leq i \leq n$, so that B is linearly independent.

1.18 Corollary: *Let $a_0, a_1, \cdots, a_\ell$ be distinct points in W. Let $P = \langle a_0, a_1, \cdots, a_\ell \rangle$. Then $\{a_0, a_1, \cdots, a_\ell\}$ is affinely independent if and only if $\dim(P) = \ell$.*

1.19 Note: For the rest of this section, we let W denote a fixed vector space over $\mathbb{R}$.

1.20 Definition: For $a, b \in W$, the **line segment** between a and b in W is the set

$$[a,b] = \{a + t(b-a) \,|\, t \in \mathbb{R}, 0 \leq t \leq 1\} = \{sa + tb \,|\, 0 \leq s, t \in \mathbb{R}, s + t = 1\}.$$

1.21 Definition: A non-empty set $\emptyset \neq C \subseteq W$ is called **convex** when it has the property that for all $a, b \in C$, we have $[a,b] \subseteq C$.

1.22 Theorem: *The intersection of a set of convex sets in W is either empty or convex.*

Proof: Let K be a set, and for each $k \in K$ let C_k be a convex set in W. Let $D = \bigcap_{k \in K} C_k$. Suppose that $D \neq \emptyset$. Let $a, b \in D$. Then $a, b \in C_k$ for all $k \in K$. Since each C_k is convex, it follows that $[a,b] \subseteq C_k$ for every $k \in K$, and so we have $[a,b] \subseteq D$.

1.23 Definition: Let $\emptyset \neq A \subseteq W$. The **convex hull** of A in W, denoted by $[A]$, is the smallest convex set in W which contains A. Equivalently, $[A]$ is the intersection of all convex sets in W which contain A.

1.24 Theorem: *Let $\emptyset \neq A \subseteq W$. Then*

$$[A] = \Big\{ \sum_{i=0}^{m} s_i a_i \Big| m \in \mathbb{N}, a_i \in A, 0 \leq s_i \in \mathbb{R}, \sum_{i=0}^{m} s_i = 1 \Big\}.$$

Proof: Let C denote the set on the right. We claim that C is convex. Let $x, y \in C$, say $x = \sum_{i=0}^{m} s_i a_i$ and $y = \sum_{i=0}^{m} t_i a_i$ where $m \in \mathbb{N}$, $0 \leq s_i, t_i$ and $\sum s_i = \sum t_i = 1$ (we can use the same upper limits, and the same points a_i, in the sums for x and y because some of the coefficients s_i, t_i can be zero). Let $z \in [x, y]$, say $z = x + r(y-x)$ where $0 \leq r \leq 1$. Then we have $z = \sum s_i a_i + r\big(\sum t_i a_i - \sum s_i a_i\big) = \sum r_i a_i$ where $r_i = s_i + r(t_i - s_i)$. Since $0 \leq s_i$ and $0 \leq t_i$ and $r_i \in [s_i, t_i]$, we must have $r_i \geq 0$. Also, we have $\sum r_i = \sum s_i + r\big(\sum t_i - \sum s_i\big) = 1 + r(1-1) = 1$, and so $z = \sum r_i a_i \in C$. Thus C is convex, as claimed. Since C is convex, and clearly $A \subseteq C$, we have $[A] \subseteq C$.

Let D be any convex set with $A \subseteq D$. For each $k \in \mathbb{N}$, let

$$C_k = \Big\{ \sum_{i=0}^{k} s_i a_i \Big| a_i \in A, 0 \leq s_i \in \mathbb{R}, \sum_{i=0}^{k} s_i = 1 \Big\}.$$

We claim that each $C_k \subseteq D$. Note that $C_0 = A \subseteq D$. Fix $k \geq 1$ and suppose, inductively, that $C_{k-1} \subseteq D$. Let $x \in C_k$, say $x = \sum_{i=0}^{k} s_i a_i$ with $0 \leq s_i$, $\sum s_i = 1$. If $s_k = 1$ then $x = a_k$ and so $x \in A \subseteq D$. Suppose that $s_k \neq 1$. Let $y = \sum_{i=0}^{k-1} \frac{s_i}{1-s_k} a_i$. Note that each $\frac{s_i}{1-s_k} \geq 0$ and that $\sum_{i=0}^{k-1} \frac{s_i}{1-s_k} = \frac{1}{1-s_k} \sum_{i=0}^{k-1} s_i = \frac{1}{1-s_k}(1-s_k) = 1$ and so we have $y \in C_{k-1} \subseteq D$. Also, we have $(1-s_k)y = \sum_{i=0}^{k-1} s_i a_i = x - s_k a_k$ and so $x = (1-s_k)y + s_k a_k = y + s_k(a_k - y) \in [y, a_k]$. Since $y \in C_{k-1} \subseteq D$ and $a_k \in D$, and $x \in [y, a_k]$, and D is convex, it follows that $x \in D$. Thus $C_k \subseteq D$. By induction, we have $C_k \subseteq D$ for all $k \in \mathbb{N}$, and hence $C = \bigcup_{k=0}^{\infty} C_k \subseteq D$. Since C is contained in every convex set D with $A \subseteq D$, it follows that $C \subseteq [A]$.

1.25 Definition: Let $\emptyset \neq A \subseteq W$. A **convex combination** on A is a point in $[A]$, that is a point of the form

$$p = \sum_{i=0}^{m} s_i a_i \text{ where } m \in \mathbb{N},\ a_i \in A,\ 0 \leq s_i \in \mathbb{R},\ \sum_{i=0}^{m} s_i = 1\,.$$

1.26 Definition: Let $\ell \in \mathbb{N}$. An (ordered, non-degenerate) **ℓ-simplex** in W is a convex set of the form $[a_0, a_1, \cdots, a_\ell]$ where $(a_0, a_1, \cdots, a_\ell)$ is an ordered $(\ell+1)$-tuple of distinct points $a_i \in W$ such that $\{a_0, a_1, \cdots, a_\ell\}$ is affinely independent. A 0-simplex is sometimes called a **point** (although it is actually a one-element set containing a point), a 1-simplex is called a **line segment**, a 2-simplex is called a **triangle**, and a 3-simplex is called a **tetrahedron**.

1.27 Definition: Let $S = [a_0, a_1, \cdots, a_\ell]$ be an ℓ-simplex in W. For each pair (j, k) with $0 \leq j < k \leq \ell$, the **medial hyperplane** $M_{j,k}$ of S is given by

$$M_{j,k} = \big\langle \tfrac{1}{2}(a_j + a_k), a_i \big| i \neq j, k \big\rangle.$$

1.28 Note: Given an ℓ-simplex $[a_0, a_1, \cdots, a_\ell]$ and a pair (j,k) with $1 \leq j < k \leq \ell$, note that the set $\left\{\frac{1}{2}(a_j + a_k), a_i \middle| i \neq j,k\right\}$ is affinely independent. Indeed if we have $s \cdot \frac{1}{2}(a_j + a_k) + \sum\limits_{i \neq j,k} s_i a_i = 0$ with $s + \sum s_i = 0$ then, letting $s_j = s_k = \frac{1}{2}s$, we have $\sum\limits_{i=0}^{\ell} s_i a_i = 0$ with $\sum s_i = 0$, and so each $s_i = 0$ (including s_j and s_k) because $\{a_0, a_1, \cdots, a_\ell\}$ is affinely independent. It follows that $\dim(M_{j,k}) = \ell - 1$. We remark that when $\ell \neq n$, the affine space $M_{j,k}$ is not a hyperplane in W but rather a hyperplane in the affine span $\langle a_0, a_1, \cdots, a_\ell \rangle$.

1.29 Theorem: *Let $[a_0, a_1, \cdots, a_\ell]$ be an ℓ-simplex in W. Then the medial hyperplanes $M_{j,k}$ have a unique point of intersection g, called the* **centroid** *of the simplex, which is given by*

$$g = \frac{1}{\ell+1} \sum_{i=0}^{\ell} a_i.$$

Proof: First we show that the point $g = \frac{1}{\ell+1} \sum\limits_{i=0}^{\ell} a_i$ lies on each medial hyperplane $M_{j,k}$. For $1 \leq j < k \leq \ell$ we have

$$g = \frac{1}{\ell+1} \sum_{i=0}^{\ell} a_i = \frac{1}{\ell+1}(a_j + a_k) + \frac{1}{\ell+1} \sum\nolimits_{i \neq j,k} a_i = \frac{2}{\ell+1} \cdot \frac{1}{2}(a_j + a_k) + \sum_{i \neq j,k} \frac{1}{\ell+1} a_i.$$

The sum of the coefficients is $\frac{2}{\ell+1} + (\ell - 1)\frac{1}{\ell+1} = \frac{\ell+1}{\ell+1} = 1$ and so $g \in M_{j,k}$.

To show that g is the unique point which lies in every medial hyperplane $M_{j,k}$, we shall show that there can be at most one point which lies in each medial hyperplane $M_{0,k}$. To do this we first show that $a_k \notin M_{0,k}$. Suppose, for a contradiction, that $a_k \in M_{0,k}$, say $a_k = s \cdot \frac{1}{2}(a_0 + a_k) + \sum\limits_{i \neq 0,k} s_i a_i$ with $s + \sum\limits_{i \neq 0,k} s_i = 1$. Then by letting $s_0 = \frac{s}{2}$ and $s_k = \frac{s}{2} - 1$ we obtain $\sum\limits_{i=0}^{\ell} s_i a_i = 0$ with $\sum s_i = 0$. Since $\{a_0, a_1, \cdots, a_\ell\}$ is affinely independent, it follows that each $s_i = 0$. But it is not possible to have both $0 = s_0 = \frac{s}{2}$ and $0 = s_k = \frac{s}{2} - 1$, so we obtain the desired contradiction. Thus $a_k \notin M_{0,k}$.

To complete the proof, we shall show that there can be at most one point which lies in each of the medial hyperplanes $M_{0,k}$. We do this by a dimension count. Let $P_k = \bigcap_{i=1}^{k} M_{0,k}$. Note that $P_k \neq \emptyset$ since we know that $g \in P_k$, and so P_k is an affine space. For $k \geq 2$ we have $P_k = P_{k-1} \cap M_{0,k}$ so that $P_k \subseteq P_{k-1}$, and we have $a_k \in P_{k-1}$ but $a_k \notin P_k$ and so $P_k \subsetneqq P_{k-1}$. Thus we have

$$M_{0,1} = P_1 \supsetneqq P_2 \supsetneqq \cdots \supsetneqq P_\ell.$$

Since $\dim(P_1) = \dim(M_{0,1}) = \ell - 1$, and since $P_k \subsetneqq P_{k-1}$ so that $\dim(P_k) < \dim P_{k-1}$ for all $k \geq 2$, we must have $\dim(P_k) \leq \ell - k$ for all k. In particular $\dim(P_\ell) \leq 0$ and hence $\dim(P_\ell) = 0$ and so P_ℓ is a one-element set containing a point, indeed $P_\ell = \{g\}$.